arXiv:1008.1703v2 [math.CO] 6 Jun 2012

THE SPLIT DECOMPOSITION OF A k-DISSIMILARITY MAP

SVEN HERRMANN AND VINCENT MOULTON

Abstract. A *k-dissimilarity map* on a finite set X is a function $D : \binom{X}{k} \to \mathbb{R}$ assigning a real value to each subset of X with cardinality k, $k \geq 2$. Such functions, also sometimes known as *k-way dissimilarities*, *k-way distances*, or *k-semimetrics*, are of interest in many areas of mathematics, computer science and classification theory, especially 2-dissimilarity maps (or *distances*) which are a generalisation of metrics. In this paper, we show how regular subdivisions of the kth hypersimplex can be used to obtain a canonical decomposition of a k-dissimilarity map into the sum of simpler k-dissimilarity maps arising from bipartitions or *splits* of X. In the special case $k = 2$, this is nothing other than the well-known *split decomposition* of a distance due to Bandelt and Dress [Adv. Math. **92** (1992), 47–105], a decomposition that is commonly to construct phylogenetic trees and networks. Furthermore, we characterise those sets of splits that may occur in the resulting decompositions of k-dissimilarity maps. As a corollary, we also give a new proof of a theorem of Pachter and Speyer [Appl. Math. Lett. **17** (2004), 615–621] for recovering k-dissimilarity maps from trees.

1. Introduction

Throughout this paper we assume $X = \{1, \ldots, n\}$, $n \geq 1$ a natural number. For $1 < k < n$, a *k-dissimilarity map* on X is a function $D : \binom{X}{k} \to \mathbb{R}$ assigning a real value to each subset of X with cardinality k (or, alternatively stated, a totally symmetric function $D : X^k \to \mathbb{R}$). Such maps are of interest in many areas of mathematics, computer science and classification theory, especially 2-dissimilarity maps (or *distances*), which are a generalisation of metrics (cf. Deza and Laurent [6]). Note that 3-dissimilarities have been investigated, for example, in [9], [17] and [10], and arbitrary k-dissimilarities in [5] and [23], under names such as *k-way dissimilarities*, *k-way distances* and *k-semimetrics*.

Here we are interested in how to decompose k-dissimilarity maps into a sum of simpler k-dissimilarity maps. Note, that various ways have been proposed to decompose distances (cf. Deza and Laurent [6]) although to our best knowledge not much is known for $k \geq 3$. More specifically, we shall introduce a generalisation of the *split decomposition* for distances that was originally introduced by Bandelt and

Date: 8th June 2012.

2010 *Mathematics Subject Classification.* 51K05, 05C05, 52B11, 92D15.

Key words and phrases. dissimilarity map, polytope, split decomposition, metric space, distance.

Dress [1]. The split decomposition is of importance in phylogenetics, where it is used to construct phylogenetic trees and networks (see e.g. Huson and Bryant [14]). Note that k-dissimilarity maps arise naturally from such trees (see e.g. Figure 1.1 and, [18, 20]); we shall discuss this connection further in Section 7.

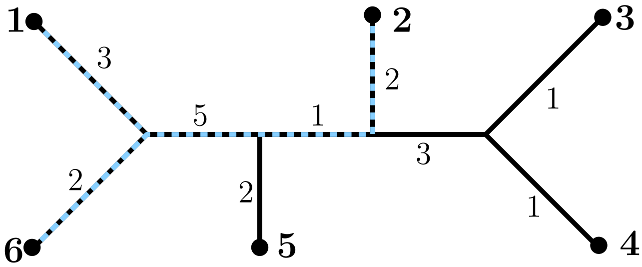


FIGURE 1.1. A weighted tree, labelled by the set $X = \{\mathbf{1}, \mathbf{2}, \dots, \mathbf{6}\}$. A k-dissimilarity map can be defined on X by assigning the length of the subtree spanned by a k-subset to that subset. For example, if $k = 3$, the subset $\{\mathbf{1}, \mathbf{2}, \mathbf{6}\}$ would be assigned the value 13.

We now explain the basic ideas underlying our results (see Section 2 for full definitions of the terminology that we use). Decompositions of k-dissimilarity maps arise in the context of polyhedral decompositions [4] as follows. Let $\Delta(k, n)$ denote the *kth hypersimplex* $\Delta(k, n) \subset \mathbb{R}^n$, that is, the convex hull of all $0/1$-vectors in $\mathbb{R}^n$ having exactly k ones. Clearly, k-dissimilarity maps on the set X are in bijection with real-valued maps from the vertices of $\Delta(k, n)$ since we can identify the vertices of $\Delta(k, n)$ with subsets of X of cardinality k. In particular, it follows that each k-dissimilarity map D gives rise to a (regular) subdivision of $\Delta(k, n)$ into smaller polytopes or *faces*. We shall call a decomposition $D = D_1 + D_2$ of D *coherent*, if the subdivisions of $\Delta(k, n)$ corresponding to D_1 and D_2 have a common refinement, which is essentially a subdivision of $\Delta(k, n)$ which contains both subdivisions.

The simplest possible regular subdivision of the polytope $\Delta(k, n)$ is a *split subdivision* (or *split* of $\Delta(k, n)$) [13], that is, a subdivision having exactly two maximal faces. As we shall show, using the polyhedral Split Decomposition Theorem [13, Theorem 3.10], it follows that a k-dissimilarity map D can always be coherently decomposed as follows. To each bipartition or *split* $S = \{A, B\}$ of X associate the *split k-dissimilarity*, defined by

$$\delta_S^k(K) := \begin{cases} 1, & \text{if } A \cap K, B \cap K \neq \emptyset, \\ 0, & \text{else,} \end{cases} \text{for all } K \in \binom{X}{k}.$$

In addition, define the *split index* α_S^D of D with respect to S in case S is non-trivial (i.e., $|A|, |B| > 1$) to be the maximal $\lambda \in \mathbb{R}_{\geq 0}$ such that $D = (D - \lambda\delta_S^k) + \lambda\delta_S^k$ is a coherent decomposition of D. If $\alpha_S^D = 0$ for all splits S of X, we call D *split-prime*. We prove the following:

Theorem 1.1 (Split Decomposition Theorem of a k-Dissimilarity Map). *Each k-dissimilarity map D on X has a coherent decomposition*

$$D = D_0 + \sum_{S \text{ split of } X} \alpha_S^D \delta_S^k , \tag{1.1}$$

where D_0 is split-prime. Moreover, this is unique among all coherent decompositions of D into a sum of split k-dissimilarities and a split-prime k-dissimilarity map.

In case D is a distance (i.e., $k = 2$) the decomposition in this theorem is precisely the split decomposition of Bandelt and Dress [1] mentioned above. For such maps, it was shown in [1, Theorem 3] that the set $\mathcal{S}_D$ of splits S with $\alpha_S^D > 0$, enjoys a special property in that it is *weakly compatible*, that is, there do not exist (pairwise distinct) $i_0, i_1, i_2, i_3 \in X$ and $S_1, S_2, S_3 \in \mathcal{S}_D$ with $S_l(i_0) = S_l(i_m)$ if and only if $m = l$, where $S(i)$ denotes the element in the split S that contains i.

In this paper we shall show that for a general k-dissimilarity D, the set $\mathcal{S}_D$ of splits with positive split index α_S^D can be characterised in a similar manner. In particular, calling any such set of splits k-*weakly compatible*, we prove the following (see Figure 1.2):

Theorem 1.2. *Let $\mathcal{S}$ be a set of splits of X. Then $\mathcal{S}$ is k-weakly compatible if and only if none of the following conditions hold:*

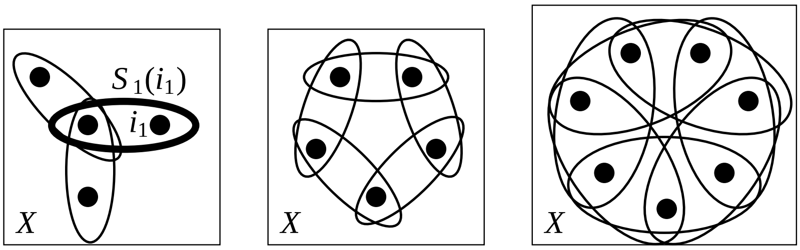
(a) There exist (pairwise distinct) $i_0, i_1, i_2, i_3 \in X$ and $S_1, S_2, S_3 \in \mathcal{S}$ with $S_l(i_0) = S_l(i_m) \iff m = l$ and $\left|X \setminus (S_1(i_0) \cup S_2(i_0) \cup S_3(i_0))\right| \geq k - 2$.

(b) For some $1 \leq \nu < k$ there exist (pairwise distinct) $i_1, \ldots, i_{2\nu+1} \in X$ and $S_1, \ldots, S_{2\nu+1} \in \mathcal{S}$ with $S_l(i_l) = S_l(i_m) \iff m \in \{l, l+1\}$ (taken modulo $2\nu + 1$) and $\left|X \setminus \bigcup_{l=1}^{2\nu+1} S_l(i_l)\right| \geq k - \nu$.

(c) For some $7 \leq \nu < 3k$ with $\nu \not\equiv 0 \mod 3$ there exist (pairwise distinct) $i_1, \ldots, i_\nu \in X$ and $S_1, \ldots, S_\nu \in \mathcal{M}$ with $S_l(i_l) = S_l(i_m) \iff m \in \{l, l+1, l+2\}$ (taken modulo ν) and $\left|X \setminus \bigcup_{l=1}^{\nu} S_l(i_l)\right| \geq k - \lfloor \nu/3 \rfloor$.

The proof of this characterisation will occupy a significant part of this paper (Section 5). Note that it immediately follows from this theorem that any k-weakly compatible set of splits is weakly compatible, since the situation pictured in Figure 1.2 (a) is the configuration that is excluded for weakly compatible sets of splits in case $k = 2$ (not including the cardinality constraint in Theorem 1.2 (a) which is always satisfied for $k = 2$). Also, in the special case where D is a k-dissimilarity map arising from a tree (as in [11]), we will further show that Theorem 1.1 can be used to recover the tree from D (see Theorem 7.2). This gives a new proof of the main theorem of Pachter and Speyer in [19].

This rest of this paper is organised as follows. We begin by presenting some definitions concerning subdivisions and splits of convex polytopes (Section 2), as well as a short discussion on splits of hypersimplices (Section 3). In Section 4, we prove Theorem 1.1, while Section 5 is devoted to the rather technical proof of



(a) (b) (c)

Figure 1.2. An illustration of the forbidden situations (a)–(c) in Theorem 1.2. The dots denote the elements $i_l \in X$ and each of the ellipses corresponds to one of the splits S_l. For example, the dots in (a) represent the elements i_0, i_1, i_2, i_3, the central dot represents the element i_0, the ellipses correspond to the splits S_1, S_2, S_3, and the dots inside the bold ellipse form the set $S_1(i_1)$. The situations in (b) and (c) correspond to the cases $\nu = 1$ and $\nu = 7$, respectively.

Theorem 1.2. This is followed by some corollaries of our main theorems related to k-weak compatibility (Section 6) and tree reconstruction (Section 7), respectively. In the last section, we present some remarks on the connection of our results with tight-spans and tropical geometry as well as some open problems.

Acknowledgements: The first author thanks the German Academic Exchange Service (DAAD) for its support through a fellowship within the Postdoc-Programme and the UEA School of Computing Sciences for hosting him during the writing of this paper.

2. Subdivisions and Splits of Convex Polytopes

We refer the reader to Ziegler [24] and De Loera, Rambau, and Santos [4] for further details concerning polytopes and subdivisions of polytopes, respectively. Let $n \geq 1$ and $P \subset \mathbb{R}^n$ be a convex polytope. For technical reasons, we assume that P has dimension $n-1$ and the origin is not an interior point of P. For any hyperplane H for which P is entirely contained in one of the two halfspaces defined by H, the intersection $P \cap H$ is called a *face* of P. A *subdivision* of P is a collection Σ of polytopes (the *faces* of Σ) such that

- ▹ $\bigcup_{F \in \Sigma} F = P$,
- ▹ for all $F \in \Sigma$ all faces of F are in Σ,
- ▹ for all $F_1, F_2 \in \Sigma$ the intersection $F_1 \cap F_2$ is a face of F_1 and F_2,
- ▹ for all $F \in \Sigma$ all vertices of F are vertices of P.

Consider a weight function $w : \operatorname{Vert} P \to \mathbb{R}$ assigning a weight to each vertex of P. This gives rise to the *lifted polytope* $\mathcal{L}_w(P) := \operatorname{conv}\left\{(v, w(v)) \in \mathbb{R}^{n+1} \;\middle|\; v \in \operatorname{Vert} P\right\}$. By projecting back to the affine hull of P, the complex of lower faces of $\mathcal{L}_w(P)$ (with respect to the last coordinate) induces a polytopal subdivision $\Sigma_w(P)$ of P. Such

a subdivision of P is called a *regular subdivision.* For two subdivisions Σ_1, Σ_2 of a polytope P, we can form the collection of polytopes

$$\Sigma := \{F_1 \cap F_2 \,|\, F_1 \in \Sigma_1, F_2 \in \Sigma_2\}\,. \tag{2.1}$$

Clearly, Σ satisfies all but the last condition for a subdivision. If this last condition is also satisfied, the subdivision Σ is called the *common refinement* of Σ_1 and Σ_2.

A *split* S of P is a subdivision of P which has exactly two maximal faces denoted by S_+ and S_- (see [13] for details on splits of polytopes). By our assumptions, the linear span of $S_+ \cap S_-$ is a linear hyperplane H_S, the *split hyperplane* of S with respect to P. Conversely, it is easily seen that a (possibly affine) hyperplane defines a split of P if and only if its intersection with the (relative) interior of P is nontrivial and it does not separate any edge of P. A set $\mathcal{T}$ of splits of P is called *compatible* if for all $S_1, S_2 \in \mathcal{T}$ the intersection of $H_{S_1} \cap H_{S_2}$ with the relative interior of P is empty. It is called *weakly compatible* if $\mathcal{T}$ has a common refinement.

Lemma 2.1. *Let P be a polytope and $\mathcal{T}$ a set of splits of P. Then $\mathcal{T}$ is weakly compatible if and only if there does not exist a set $\mathcal{H} \subset \{H_S \,|\, S \in \mathcal{T}\}$ of splitting hyperplanes and a face F of P such that $F \cap \bigcap_{H \in \mathcal{H}} H = \{x\}$ and x is not a vertex of P.*

Proof. Obviously, if there is a set of hyperplanes $\mathcal{H} \subset \{H_S \,|\, S \in \mathcal{T}\}$ with this property, the set $\mathcal{T}$ cannot have a common refinement and hence is not compatible. Conversely, we can iteratively compute the collections (2.1) for elements of $\mathcal{T}$ and it has to happen at some stage that there occurs an additional vertex v. At this stage take F to be the minimal face of P containing v and $\mathcal{H} = \{H_S \,|\, v \in H_S, S \in \mathcal{T}\}$. □

For a split S, it is easy to explicitly define a weight function w_S such that $S = \Sigma_{w_S}(P)$, hence all splits of P are regular subdivisions of P; see [13, Lemma 3.5]. Finally, as mentioned in the introduction, a sum $w = w_1 + w_2$ of two weight functions for P is called *coherent* if $\Sigma_w(P)$ is the common refinement of $\Sigma_{w_1}(P)$ and $\Sigma_{w_2}(P)$. So a sum $\sum_{S \in \mathcal{T}} \lambda_S w_S$ with $\lambda_S \in \mathbb{R}_{>0}$ is coherent if and only if the set $\mathcal{T}$ of splits is weakly compatible.

3. Splits of Hypersimplices

Let $n > k > 0$. As mentioned above, the *kth hypersimplex* $\Delta(k,n) \subset \mathbb{R}^n$ is defined as the convex hull of all $0/1$-vectors in $\mathbb{R}^n$ having exactly k ones, or, equivalently, $\Delta(k,n) = [0,1]^n \cap \{x \in \mathbb{R}^n \,|\, \sum_{i=1}^n x_i = k\}$. The polytope $\Delta(k,n)$ is $(n-1)$-dimensional and has $2n$ facets defined by $x_i = 1, x_i = 0$ for $1 \le i \le n$. Each face of $\Delta(k,n)$ is isomorphic to $\Delta(k', n')$ for some $k' \le k$, $n' < n$. This polytope first appeared in the work of Gabriélov, Gel′fand and Losik [8, Section 1.6].

For a split $\{A, B\}$ of X, and $\mu \in \mathbb{N}$ the *(A, B, μ)-hyperplane* is defined by the equation

$$\mu \sum_{i \in A} x_i \;=\; (k - \mu) \sum_{i \in B} x_i\,. \tag{3.1}$$

The splits of $\Delta(k,n)$ can then be characterised as follows:

Proposition 3.1 (Lemma 5.1 and Proposition 5.2 in [13])**.** *The splits of* $\Delta(k,n)$ *are given by the* (A,B,μ)*-hyperplanes with* $k-\mu+1 \le |A| \le n-\mu-1$ *and* $1 \le \mu \le k-1$*.*

We will be interested in the special class of splits of $\Delta(k,n)$ defined by subsets of X. For $A \subsetneq X$ define the hyperplane $H_A \subset \mathbb{R}^n$ by

$$\sum_{i\in A} x_i = 1\,. \tag{3.2}$$

Corollary 3.2. *For* $A \subset X$ *the hyperplane* H_A *defines a split of* $\Delta(k,n)$ *if and only if* $2 \le |A| \le n-k$*. Otherwise,* H_A *defines the trivial subdivision of* $\Delta(k,n)$*.*

Proof. Since $\sum_{i=1}^n x_i = k$ for all $x \in \Delta(k,n)$, the hyperplane H_A defines the same split as the $(X \setminus A, A, 1)$-hyperplane. Thus, by Proposition 3.1, H_A defines a split if and only if $k \le n-|A| \le n-2$, which is equivalent to $2 \le |A| \le n-k$. Obviously, if $|A| \le 1$ or $|A| > k$, the hyperplane H_A does not meet the interior of $\Delta(k,n)$ hence defines the trivial subdivision. □

The split of $\Delta(k,n)$ defined by H_A for some $A \subset X$ will be called S_A. We now characterise when such splits of $\Delta(k,n)$ are compatible.

Lemma 3.3. *Let* $A,B \subset X$*. The two splits* S_A *and* S_B *of* $\Delta(k,n)$ *are compatible if and only if either* $A \subset B$*,* $B \subset A$*,* $|A \cup B| \ge n-k+2$*, or* $k=2$ *and* $A \cap B = \emptyset$*.*

Proof. By [13, Proposition 5.4], two splits of $\Delta(k,n)$ defined by $(A,B;\mu)$- and $(C,D;\nu)$-hyperplanes are compatible if and only if one of the following holds:

$$\begin{aligned} |A\cap C| &\le k-\mu-\nu\,, & \qquad |A\cap D| &\le \nu-\mu\,,\\ |B\cap C| &\le \mu-\nu\,, & \text{or}\qquad |B\cap D| &\le \mu+\nu-k\,. \end{aligned}$$

That is, the two splits S_A (defined by the $(X\setminus A, A, 1)$-hyperplane) and S_B (defined by the $(X\setminus B, B, 1)$-hyperplane) are compatible if and only if

$$\begin{aligned} |(X\setminus A)\cap(X\setminus B)| &\le k-2\,, & \qquad |(X\setminus A)\cap B| &\le 0\,,\\ |A\cap(X\setminus B)| &\le 0\,, & \text{or}\qquad |A\cap B| &\le 2-k\,. \end{aligned}$$

The first condition can be rewritten as $|A \cup B| \ge n-k+2$, the second condition is equivalent to $B \subset A$, the third condition is equivalent to $A \subset B$, and the last condition can only be true if $k=2$ and $A \cap B = \emptyset$. □

For a weight function w and a split S_A of $\Delta(k,n)$, we define the *split index* $\alpha^w_{S_A}$ of w with respect to S_A as

$$\alpha^w_{S_A} = \max\left\{\lambda \in \mathbb{R}_{\ge 0} \,\middle|\, (w-\lambda w_{S_A}) + \lambda w_{S_A} \text{ is coherent}\right\},$$

where w_{S_A} is a weight function inducing the split S_A on $\Delta(k,n)$. Note, that this is the coherency index of the weight function w with respect to w_{S_A} as defined in [13, Section 2].

4. The Split Decomposition of a k-Dissimilarity Map

In this section, we shall prove Theorem 1.1. We begin with some preliminaries concerning the relationship between splits of X and splits of $\Delta(k,n)$.

As mentioned in the introduction, we can identify vertices of $\Delta(k,n)$ with subsets of X of cardinality k. With this identification in mind, for a k-dissimilarity map D, define the weight function $w_D : \operatorname{Vert}\Delta(k,n) \to \mathbb{R};\ K \mapsto -D(K)$ on the vertices of $\Delta(k,n)$. In addition, for $D = \delta_S^k$, we put $w_S^k := w^{\delta_S^k}$. This allows us to relate splits of X with splits of $\Delta(k,n)$.

Lemma 4.1. *Let $S = \{A, B\}$ be a non-trivial split of X.*

(a) The subdivision $\Sigma_{w_S^k}(\Delta(k,n))$ is the common refinement of the subdivisions induced on $\Delta(k,n)$ by H_A and H_B.

(b) (i) If $\min(|A|,|B|) \geq k$ then the subdivision $\Sigma_{w_S^k}(\Delta(k,n))$ is the common refinement of the splits S_A and S_B.

(ii) If $|A| < k \leq |B|$ then the subdivision $\Sigma_{w_S^k}(\Delta(k,n))$ is the split S_B.

(iii) If $\max(|A|,|B|) < k$ then the subdivision $\Sigma_{w_S^k}(\Delta(k,n))$ is trivial.

Proof. (a) By [13, Lemma 3.5], a weight function for the split S_B defined by the $(A, B, 1)$-hyperplane is given by

$$w_1(v) = \begin{cases} |\sum_{i=1}^n a_i v_i|, & \text{if } |\sum_{i=1}^n a_i v_i| > 0, \\ 0, & \text{else,} \end{cases}$$

where a is the normal vector of the $(A, B, 1)$-hyperplane. Since $\sum_{i=1}^n x_i = k$ for all $x \in \Delta(k,n)$, we have $|\sum_{i=1}^n a_i x_i| = |A \cap K| - (k-1)|B \cap K| = k(1 - |B \cap K|)$, hence (again identifying vertices of $\Delta(k,n)$ with k-subsets of X)

$$w_1(K) = \begin{cases} k, & \text{if } B \cap K = \emptyset, \\ 0, & \text{else.} \end{cases}$$

Similarly, a weight function for the split S_A is given by

$$w_2(K) = \begin{cases} k, & \text{if } A \cap K = \emptyset, \\ 0, & \text{else.} \end{cases}$$

Obviously, $\tilde{w} := \frac{w_1 + w_2}{k} + 1$ defines the same subdivision as $w_1 + w_2$, and we have $\tilde{w} = -\delta_S^k$.

(b) Follows from (a) using Corollary 3.2 and Lemma 3.3.

□

In particular, it follows from Lemma 4.1 that if $|X| \geq 2k - 1$ the subdivision $\Sigma_{w_S^k}(\Delta(k,n))$ of $\Delta(k,n)$ is not trivial for any split S, which implies in this case that the split S of X can be recovered from the subdivision $\Sigma_{w_S^k}(\Delta(k,n))$.

Furthermore, Lemma 4.1 implies that the split index α_S^D of a k-dissimilarity map D on X with respect to a non-trivial split $S = \{A, B\}$ of X can be written in terms of split indices for splits of the hypersimplex $\Delta(k,n)$ as

$$\alpha_S^D = \min(\alpha_{S_B}^{w_D}, \alpha_{S_A}^{w_D}) .$$

If $\alpha_S^D = 0$ for all non-trivial splits of X, we call D *free of non-trivial splits*. This enables us to deduce our split decomposition theorem for k-dissimilarities by using the polyhedral split decomposition theorem for weight functions. However, since our correspondence only works for non-trivial splits, we have to deal with the trivial splits as a special case before we can give our proof.

4.1. **The Trivial Splits.** Each $a \in A$ defines a trivial split $S_a := \{\{a\}, X \setminus \{a\}\}$ separating a from the rest of X. The corresponding k-dissimilarity map $\delta_{S_a}^k$ on X is given by

$$\delta_{S_a}^k(K) := \begin{cases} 1, & \text{if } a \in K, \\ 0, & \text{else.} \end{cases}$$

Hence the extension of the weight function $w_{S_a}^k = -\delta_{S_a}^k : \operatorname{Vert}\Delta(k,n) \to \mathbb{R}$ to $\mathbb{R}^n$ is linear and thus induces the trivial subdivision into $\Delta(k,n)$. In fact, $\{w_{S_a}^k \mid a \in X\}$ is a basis for the space of all functions from $\mathbb{R}^n$ to $\mathbb{R}$. This implies that $\alpha_S^{\delta_{S_a}^k} = 0$ for all $a \in X$ and all non-trivial splits S of X, so adding or subtracting k-dissimilarities corresponding to trivial splits does not interfere with split indices for non-trivial splits.

For some $a \in X$ and a k-dissimilarity map D that is free of non-trivial splits, we define the *split index* of the trivial split S_a as

$$\alpha_{S_a}^D := \frac{1}{2} \min \left\{ \min_{b,c \in X \setminus (L \cup \{a\})} (D(L,a,b) + D(L,a,c) - D(L,b,c)) \;\middle|\; L \in \binom{X \setminus \{a\}}{k-2} \right\} .$$

For an arbitrary k-dissimilarity map D we then set $\alpha_{S_a}^D := \alpha_{S_a}^{D_0}$ where D_0 is defined as

$$D_0 := D - \sum_{S \text{ non-trivial split of } X} \alpha_S^D \delta_S^k .$$

The following lemma shows that we can iteratively compute all the trivial split indices.

Lemma 4.2. *Let D be a k-dissimilarity map on X, $a, a' \in X$ distinct, and $\lambda \in \mathbb{R}_{\geq 0}$. Then*

$$\alpha_{S_a}^D = \alpha_{S_a}^{D + \lambda \delta_{S_{a'}}^k} .$$

Proof. For all $L \in \binom{X\setminus\{a\}}{k-2}$ and $b, c \in X \setminus (L \cup \{a\})$, we see that

$$\delta^k_{S_{a'}}(L,a,b) + \delta^k_{S_{a'}}(L,a,c) - \delta^k_{S_{a'}}(L,b,c) = \begin{cases} 1-1, & \text{if } a' \in L \cup \{b,c\}, \\ 0, & \text{else,} \end{cases}$$
$$= 0,$$

and hence $(D+\lambda\delta^k_{S_{a'}})(L,a,b) + (D+\lambda\delta^k_{S_{a'}})(L,a,c) - (D+\lambda\delta^k_{S_{a'}})(L,b,c) = D(L,a,b) + D(L,a,c) - D(L,b,c)$. □

4.2. **Proof of the Split Decomposition Theorem 1.1.** Recall that a k-dissimilarity map D on X is called *split-prime* if for all (trivial and non-trivial) splits S of X we have $\alpha^D_S = 0$.

Proof. Using the Split Decomposition Theorem for polytopes [13, Theorem 3.10], we obtain the decomposition

$$w_D \;=\; w_0 + \sum_{\Sigma \text{ split of } \Delta(k,n)} \alpha^{w_D}_\Sigma w_\Sigma \,,$$

of w_D, where w_Σ is a weight function defining the split Σ of $\Delta(k,n)$. Setting

$$D_0 := -\left(w_0 + \sum \alpha^{w_D}_\Sigma w_\Sigma + \sum_{A\subset X, |A|\geq 2} \left(\alpha^{w_D}_{S_A} - \alpha^D_{\{A, X\setminus A\}}\right) w_{S_A} \right),$$

where the first sum ranges over all splits Σ of $\Delta(k,n)$ that are not of the form S_A for some $A \subset X$, we can rewrite the above decomposition of D as

$$D \;= D_0 + \sum_{S \text{ non-trivial split of } X} \alpha^D_S D^k_S \,.$$

This decomposition is unique because of the uniqueness of the decomposition of w_D.

Now for all $a \in X$ we compute the split indices $\alpha^D_{S_a} = \alpha^{D_0}_{S_a}$ to derive the final split decomposition, which is again unique by Lemma 4.2. □

For a k-dissimilarity map D on X, we define $\mathcal{S}_D := \{S \text{ split of } X \,|\, \alpha^D_S \neq 0\}$, that is the set of all splits of X that appear in the Split Decomposition (1.1) and recall from the introduction that such a set is by definition k-weakly compatible.

Proposition 4.3. *A set $\mathcal{S}$ of splits of X is k-weakly compatible if and only if the set $\mathcal{T} = \{S_A \text{ split of } \Delta(k,n) \,|\, A \in S, S \in \mathcal{S}\}$ of splits of $\Delta(k,n)$ is weakly compatible.*

Proof. It follows from the Split Decomposition Theorem for polytopes [13, Theorem 3.10] that a set of splits of $\Delta(k,n)$ is weakly compatible if and only if it occurs in the split decomposition of some weight function of $\Delta(k,n)$. This implies that a set $\mathcal{S}$ of non-trivial splits is k-weakly compatible if and only if $\mathcal{T}$ is a weakly compatible set of splits of $\Delta(k,n)$. By definition, adding trivial splits does not change the k-weakly compatibility of a set, so the claim follows. □

5. Weak compatibility of $\Delta(k,n)$-splits

In this section, we prove a theorem from which Theorem 1.2 immediately follows by Proposition 4.3. For a family $\mathcal{M}$ of subsets of X, we denote by $\mathcal{T}(\mathcal{M}) := \{S_A \text{ split of } \Delta(k,n) \,|\, A \in \mathcal{M}\}$ the corresponding set of splits of $\Delta(k,n)$.

Theorem 5.1. *Let $\mathcal{M}$ be a collection of subsets of a set X. Then the set $\mathcal{T}(\mathcal{M})$ of splits of $\Delta(k,n)$ is weakly compatible if and only if none of the following conditions hold:*

(a) There exist (pairwise distinct) $i_0, i_1, i_2, i_3 \in X$ and $A_1, A_2, A_3 \in \mathcal{M}$ with $i_m \in A_l \iff m \in \{0, l\}$ and $\left|X \setminus (A_1 \cup A_2 \cup A_3)\right| \geq k-2$.

(b) For some $1 \leq \nu < k$ there exist (pairwise distinct) $i_1, \dots, i_{2\nu+1} \in X$ and $A_1, \dots, A_{2\nu+1} \in \mathcal{M}$ with $i_m \in A_l \iff (m \in \{l, l+1\}$ (taken modulo $2\nu+1$) and $\left|X \setminus \bigcup_{i=1}^{2\nu+1} A_i\right| \geq k-\nu$.

(c) For some $7 \leq \nu < 3k$ with $\nu \bmod 3 \neq 0$ there exist (pairwise distinct) $i_1, \dots, i_\nu \in X$ and $A_1, \dots, A_\nu \in \mathcal{M}$ with $i_m \in A_l \iff m \in \{l, l+1, l+2\}$ (taken modulo ν) and $\left|X \setminus \bigcup_{i=1}^{\nu} A_i\right| \geq k - \lfloor \nu/3 \rfloor$.

5.1. **Sufficiency of Conditions (a)–(c).** (a): Suppose (a) holds. Choose a subset B of $X \setminus (A_1 \cup A_2 \cup A_3)$ with $|B| = k-2$ and consider the face F of $\Delta(k,n)$ defined by the facets $x_i = 1$ for $i \in B$ and $x_i = 0$ for $i \in X \setminus (B \cup \{i_0, i_1, i_2, i_3\})$. Looking at the intersection $I := F \cap H_{A_1} \cap H_{A_2} \cap H_{A_3}$ we have

$$x_{i_0} + x_{i_1} = x_{i_0} + x_{i_2} = x_{i_0} + x_{i_3} = 1 \text{ and } x_{i_0} + x_{i_1} + x_{i_2} + x_{i_3} = 2 \text{ for all } x \in I\,.$$

This yields $x_{i_k} = 1 - x_{i_0}$ for $k \in \{1,2,3\}$ and eventually $x_{i_k} = 1/2$ for all $k \in \{0,1,2,3\}$. Hence we have $I = \{x\}$ where $x \in \mathbb{R}^n$ is defined via

$$x_i = \begin{cases} 1, & \text{if } i \in B, \\ \frac{1}{2}, & \text{if } i \in \{i_0, i_1, i_2, i_3\}, \\ 0, & \text{else.} \end{cases}$$

By Lemma 2.1, $\mathcal{T}(\mathcal{M})$ is not weakly compatible.

(b): Suppose (b) holds. Choose a subset B of $X \setminus \bigcup_{i=1}^{2\nu+1} A_i$ with $|B| = k-\nu$ together with some $m \in B$ and consider the face F of $\Delta(k,n)$ defined by the facets $x_i = 1$ for $i \in B \setminus \{m\}$ and $x_i = 0$ for $i \in X \setminus (B \cup \{i_1, \dots, i_{2\nu+1}\})$. We consider the intersection $I := F \cap \bigcap_{i=1}^{2\nu+1} H_{A_i}$ and get $x_{i_l} + x_{i_{l+1}} = 1$ for all $x \in I$ and $1 \leq l \leq 2\nu$. So $x_{i_l} = x_{i_{l+2}}$ for all $1 \leq l \leq 2\nu - 1$ which implies $x_{i_1} = x_{i_{2\nu+1}}$ and, since $x_{i_{2\nu+1}} + x_{i_1} = 1$, we have $x_{i_l} = 1/2$ for all $1 \leq k \leq 2\nu+1$. Since $\sum_{i=1}^{2\nu+1} x_i + x_m = \nu$ we also get $x_m = 1/2$. Hence, we have $I = \{x\}$ where $x \in \mathbb{R}^n$ is defined via

$$x_i = \begin{cases} 1, & \text{if } i \in B \setminus \{m\}, \\ \frac{1}{2}, & \text{if } i \in \{i_1, \dots i_{2\nu+1}, m\}, \\ 0, & \text{else.} \end{cases}$$

By Lemma 2.1, $\mathcal{T}(\mathcal{M})$ is not weakly compatible.

(c): Suppose (c) holds. Choose a subset B of $X \setminus \bigcup_{i=1}^{\nu} A_i$ with $|B| = k - \lfloor \nu/3 \rfloor$ together with some $m \in B$ and consider the face F of $\Delta(k, n)$ defined by the facets $x_i = 1$ for $i \in B \setminus \{m\}$ and $x_i = 0$ for $i \in X \setminus (B \cup \{i_1, \ldots, i_\nu\})$. We consider the intersection $I := F \cap \bigcap_{i=1}^{\nu} H_{A_i}$ and get $x_{i_l} + x_{i_{l+1}} + x_{i_{l+2}} = 1$ for all $x \in I$ and $1 \leq l \leq \nu$. As in Case (b) we obtain $x_{i_l} = 1/3$ for all $1 \leq k \leq \nu$ and, since $\sum_{i=1}^{\nu} x_i + x_m = \lfloor \nu/3 \rfloor$, we get $x_m = \overline{\nu}/2$, where $\overline{\nu} = \nu \bmod 3$. Hence, we have $I = \{x\}$ where $x \in \mathbb{R}^n$ is defined via

$$x_i = \begin{cases} 1, & \text{if } i \in B \setminus \{m\}, \\ \frac{1}{3}, & \text{if } i \in \{i_1, \ldots i_\nu, m\}, \\ \frac{\overline{\nu}}{3}, & \text{if } i = m, \\ 0, & \text{else.} \end{cases}$$

By Lemma 2.1, $\mathcal{T}(\mathcal{M})$ is not weakly compatible. □

5.2. **Necessity of Conditions (a)–(c).** Suppose $\mathcal{T}(\mathcal{M})$ is not weakly compatible and that none of (a) – (c) hold. Then, by Lemma 2.1, there exists some subset $\mathcal{M}' \subset \mathcal{M}$ and some face F of $\Delta(k, n)$ such that $I := F \cap \bigcap_{A \in \mathcal{M}'} H_A = \{x\}$, x not a vertex of $\Delta(k, n)$. We assume that $\mathcal{M}'$ is minimal with this property and denote by $X' \subset X$ the set of coordinates not fixed to 0 or 1 in F, that is, $0 < x_i < 1$ if and only if $i \in X'$. For any $i \in X'$ we denote by $\mathcal{M}(i) := \{A \in \mathcal{M}' \mid i \in A\}$ the set of all $A \in \mathcal{M}'$ containing i.

We first state some simple facts for later use:

(F1) For all distinct $i, j \in X'$, we have $\mathcal{M}(i) \neq \mathcal{M}(j)$.
(F2) For all distinct $A, B \in \mathcal{M}'$, we have $A \not\subset B$.
(F3) For all $A \in \mathcal{M}'$, we have $|A \cap X'| \geq 2$.
(F4) For all $A \in \mathcal{M}'$, there exists some $i \in A$ with $|\mathcal{M}(i)| \geq 2$.

Proof. (F1) Suppose there exist distinct $i, j \in X'$, with $\mathcal{M}(i) = \mathcal{M}(j)$. Then choose some $0 < \epsilon < \min(x_i, 1 - x_j)$ and consider $x' \in \mathbb{R}^n$ defined by

$$x'_l = \begin{cases} x_l - \epsilon, & \text{if } l = i, \\ x_l + \epsilon, & \text{if } l = j, \\ x_l, & \text{else.} \end{cases}$$

So $x \neq x'$ and $x' \in I$, a contradiction.

(F2) Follows from the minimality of $\mathcal{M}'$.
(F3) Suppose $|A \cap X'| = \{j\}$ for some $A \in \mathcal{M}'$ and $j \in X'$. Then $0 < x_j < 1$ but $x_i \in \{0, 1\}$ for all $i \in A \setminus \{j\}$ which obviously contradicts $\sum_{i \in A} x_i = 1$.
(F4) Let $A \in \mathcal{M}'$. By F3 there exist distinct $i, = j \in A$ and by F1 $\mathcal{M}(i) \neq \mathcal{M}(j)$. However, $A \in \mathcal{M}(i) \cap \mathcal{M}(j)$ so either $\mathcal{M}(i)$ or $\mathcal{M}(j)$ has to contain another $B \in \mathcal{M}'$.

□

As the next step, we will show that none of the following conditions may be satisfied:

(i) There exists (pairwise distinct) $i_0, i_1, i_2, i_3 \in X'$ and $A_1, A_2, A_3 \in \mathcal{M}'$ with $i_m \in A_l \iff m \in \{0, l\}$.
(ii) For some $\nu \in \mathbb{N}$, there exist (pairwise distinct) $i_1, \ldots, i_{2\nu+1} \in X'$ and $A_1, \ldots, A_{2\nu+1} \in \mathcal{M}'$ with $i_m \in A_l \iff m \in \{l, l+1\}$ (taken modulo $2\nu + 1$).
(iii) For some $\nu \in \mathbb{N}$, there exist (pairwise distinct) $i_0, i_1, \ldots, i_{2\nu+1} \in X'$ and $A_1, \ldots, A_{2\nu+1} \in \mathcal{M}'$ with $\mathcal{M}(i_0) = \{A_1\}, \mathcal{M}(i_{2\nu+1}) = \{A_{2\nu+1}\}$ and $\mathcal{M}(i_l) = \{A_l, A_{l+1}\}$ for $1 \le l \le 2\nu$.
(iv) For some $\nu \in \mathbb{N}$, there exist (pairwise distinct) $i_1, \ldots, i_{2\nu} \in X'$ and $A_1, \ldots, A_{2\nu} \in \mathcal{M}'$ with $\mathcal{M}(i_l) = \{A_l, A_{l+1}\}$ (taken modulo 2ν).
(v) There exists some $i \in X'$ with $|\mathcal{M}(i)| = 3$.
(vi) For some $A \in \mathcal{M}'$, there exist distinct $i, j \in A$ such that $|\mathcal{M}(i)|, |\mathcal{M}(j)| \ge 4$.

Proof. (i): Suppose this were true. Then we have $\sum_{i \in A_l \setminus \{i_0\}} x_i = 1 - x_{i_0}$ for $l \in \{1, 2, 3\}$, hence $\sum_{i \in A_1 \cup A_2 \cup A_3} x_i \le x_{i_0} + \sum_{l=1}^{3} \sum_{i \in A_l \setminus \{i_0\}} x_i \le 3 - 2x_{i_0} < 3$. Since $\sum_{i \in X} x_i = k$, this implies $\sum_{i \in X \setminus (A_1 \cup A_2 \cup A_3)} x_i > k - 3$ and, because $x_i \in \{0, 1\}$ for all $i \in X \setminus (A_1 \cup A_2 \cup A_3)$, we get $|X \setminus (A_1 \cup A_2 \cup A_3)| \ge k - 2$. So we are in situation (a) of the theorem, a contradiction.

(ii): For the purpose of this proof, a collection of i_l and A_l satisfying this condition will be called a *cycle*. We set $T = \bigcup_{i=1}^{2\nu+1} A_i$, $T_1 := \{i_l \mid 1 \le l \le 2\nu + 1\}$, $T_2 := T \setminus T_1$, $t := |T|$, $t_1 := |T_1|$, and $t_2 := |T_2|$. Cycles are partially ordered by the lexicographic ordering of the pair (ν, t). We assume without loss of generality that our cycle is minimal in the set of all cycles occurring in $\mathcal{M}'$.

As base case we consider $\nu = 1$ and $t \le 5$. Each decreasing chain of cycles will eventually reach this case since $\nu \ge 1$ and $t \ge 2\nu + 1$. Then (after a possible exchange of A_3 with A_1 or A_2) we can assume that $T \subset A_1 \cup A_2$, hence $\sum_{i \in T} x_i < 2$. This implies that $\sum_{i \in X \setminus T} x_i > k - 2$ and hence $n - t \ge k - 1$ since $x_i \in \{0, 1\}$ for all $i \in X \setminus T$. So we are in situation (b) of the theorem, a contradiction.

We say that a set $A \in \mathcal{M}'$ is of *a-type* (with respect to some cycle Z) if for some $1 \le l \le 2\nu + 1$ we have $i_l \in A$, $A \subset A_l \cup A_{l+1}$, and $|A \cap T_2| \ge 2$. The set is of *b-type* (with respect to some cycle Z) if there exists some $i \in A \cap T_2$ and some $j \in A \cap (X' \setminus T)$. We will show that for the cycle Z each set A (distinct from all A_l) with $A \cap T \neq \emptyset$ is either of a-type or of b-type with respect to Z.

First consider some set A (distinct from all A_l) with $i_l \in A$ for some $1 \le l \le 2\nu+1$ and some $j \in A \setminus \{i_l\}$. Then $j \in T$ because otherwise i_l, i_{l-1}, i_{l+1}, j and A_l, A_{l+1}, A would satisfy Condition (i) for some $j \in A \setminus T$. Furthermore, if there exists some $m \notin \{l, l+1\}$ with $j \in A_m$, then we could form a smaller cycle. We get $j \in A_l \cup A_{l+1}$ and (using F2) $|A \cap T_2| \ge 2$, so A is of a-type.

Now fix a minimal cycle Z and consider an arbitrary set B (distinct from all A_l) with $B \cap T \neq \emptyset$. Suppose that $B \subset T_2$. This implies that there either exists a smaller cycle, or we have the situation that there exists some $1 \le l \le 2\nu + 1$ such that $B \subset A_{l+1} \cup A_{l-1}$ and $B \cap A_{l+1}, B \cap A_{l-1} \neq \emptyset$. By the minimality of our cycle this implies $A_l \subset T_1$. However, this implies $B \cup A_l \subsetneq A_{l+1} \cup A_{l-1}$, a contradiction

to $\sum_{i \in A} x_i = 1$ for all $B \in \mathcal{M}'$ and $x_i > 0$ for all $i \in X'$. So B either contains some element of T_1 implying B is of a-type or some element from $X \setminus T$ implying B is of b-type.

Now each $i \in T_1$ cannot be contained in some set of b-type by definition and can be contained in at most one set of a-type by F2. Furthermore, each $i \in T_2$ can be contained in at most two sets of a-type or in at most one set of b-type but not both. To see this assume that $i \in A_l$ is contained in two sets A, B either A of a-type and B of b-type or both of b-type. Then there exist $i_1 \in A \setminus (B \cup A_l)$, $i_2 \in B \setminus (A \cup A_l)$, and $i_3 \in A_l \setminus (B \cup A)$ such that A, B, A_l and i, i_1, i_2, i_3 satisfy Condition (i). For the same reason, each $i \in X' \setminus T$ can be in at most two sets of b-type.

We denote the number of sets of a-type (b-type) with respect to Z by a (by b). In order to uniquely define all t coordinates of x_i with $i \in T$, it is necessary to have at least t equations involving some x_i with $i \in T$, that is, t sets in $\mathcal{M}'$ which contain elements of T. By our considerations above, all such sets have to be either of a-type or of b-type or be equal to some A_l for $1 \leq l \leq 2\nu + 1$. Hence we get $a + b + 2\nu + 1 \geq t$, or, equivalently (since $t = t_1 + t_2 = t_2 + 2\nu + 2$),

$$a + b \geq t_2 \,. \tag{5.1}$$

Furthermore, by the fact that some $j \in T_2$ can only be in one set of b-type and this holds only if it is not in some set of a-type, we have $b \leq t_2 - a'$, where a' is the number of elements of T_2 contained in some set of a-type. Together with Inequality (5.1) we obtain

$$t_2 - a \leq b \leq t_2 - a' \,; \tag{5.2}$$

in particular $a' \leq a$. However, since each set of a-type contains at least two elements of T_2 and each element of T_2 is contained in at most two sets of a-type, which implies $a' \geq a$, we have $a' = a$ and each element of T_2 is contained in either one set of b-type or each element of T_2 is contained in exactly two sets of a-type. In view of the definition of the sets of a-type the former implies that there are no sets of a-type at all and the latter implies that the sets of a-type with respect to Z form themselves a cycle Z' together with the elements $j_1, \ldots, j_{2\nu+1} \in T_2$ contained in sets of a-type with respect to Z.

We first consider the latter case. Suppose without loss of generality that $j_l \in A_l$ and call the set of a-type containing j_l and j_{l+1} B_l. Then the sets A_l are sets of a-type with respect to Z'. Hence $\bigcup_{l=1}^{2\nu+1}(A_l \cup B_l) = \{i_l, j_l \mid 1 \leq l \leq 2\nu + 1\}$. If now $2\nu + 1$ is not divisible by 3, then we are in the situation (c) of the theorem, a contradiction, since $\nu \geq 6$ by our base case and $\nu < 3k$ obviously holds. If $2\nu + 1$ is divisible by 3, then choose some $0 < \epsilon < \min\{x_{i_l}, x_{j_l} \mid 1 \leq l \leq 2\nu + 1\}$ and consider $x' \in \mathbb{R}^n$ defined by

$$x'_l = \begin{cases} x_l + \epsilon, & \text{if } l = i_m \text{ and } m \equiv 1 \bmod 3 \text{ or } l = j_m \text{ and } m \equiv 2 \bmod 3, \\ x_l - \epsilon, & \text{if } l = j_m \text{ and } m \equiv 1 \bmod 3 \text{ or } l = i_m \text{ and } m \equiv 3 \bmod 3, \\ x_l, & \text{else.} \end{cases}$$

Then $x \neq x'$ and $x \in I$, a contradiction.

The case remaining is $a = 0$. Then Inequality (5.2) implies $b = t_2$. So $\sum_{i\in T} x_i = 2\nu + 1 - \sum_{i\in T_1} x_i$, since each element of T_2 is in exactly one of the $2\nu+1$ sets A_l and each element of T_1 in exactly two. This is equivalent to $\sum_{i\in T_1} x_i = \nu - \frac{1}{2}(\sum_{i\in T_2} x_i - 1)$. Define T_3 to be the set of all elements of X' that are one set of b-type but not in T. There cannot be any elements of X' that are in more than two sets of b-type but not in T because this would satisfy Condition (i). For some $t \in T_3$ which is in exactly one set of b-type, we get $x_t \leq 1 - x_j \leq 1 - 1/2x_j$ for some $j \in T_2$, and for some $t \in T_3$ which is in exactly two sets of b-type, we get $x_t \leq 1 - \max(x_j, x_l) \leq 1 - 1/2(x_j + x_l)$ for some $j, l \in T_2$. Since each $j \in T_2$ is contained in exactly one set of b-type, each $j \in T_2$ occurs exactly once, hence we get $\sum_{i\in T_3} x_i \leq |T_3| - \frac{1}{2}\sum_{i\in T_2} x_i$. So

$$\begin{aligned}\sum_{i\in T\cup T_3} x_i &\leq \nu - \frac{1}{2}\left(\sum_{i\in T_2} x_i - 1\right) + \sum_{i\in T_2} x_i + |T_3| - \frac{1}{2}\sum_{i\in T_2} x_i \\ &= \nu + |T_3| + \frac{1}{2},\end{aligned}$$

and $|X \setminus (T \cup T_3)| \geq k - \nu - |T_3| - 1/2)$. Hence $|X \setminus T| \geq k - \nu$ (as it has to be an integer). So we are in the situation (b) of the theorem, a contradiction

(iii),(iv): Choose some $0 < \epsilon < \min\{x_{i_l} \mid l \text{ odd}\} \cup \{1 - x_{i_l} \mid l \text{ even}\}$ and define the point $x' \in \mathbb{R}^n$ by

$$x'_l = \begin{cases} x_l - \epsilon, & \text{if } l = i_j \text{ for some odd } j, \\ x_l + \epsilon, & \text{if } l = i_j \text{ for some even } j, \\ x_l, & \text{else.}\end{cases}$$

Obviously, $x \neq x'$ and it is easily checked that $x' \in I$, a contradiction.

(v): Suppose there exists some $i \in X'$ with $\mathcal{M}(i) \geq 3$. Since Condition (i) cannot hold, there has to exist some $B \in \mathcal{M}(i)$ such that, for each $j \in B$, there exists some $B \neq C$ with $j \in C \in \mathcal{M}(i)$. By F2, there exist distinct $j_1, j_2 \in B$ and $C_1, C_2 \in \mathcal{M}(i)$ with $j_1 \in C_1$, $j_2 \in C_2$ and $l_1 \in C_1 \setminus B$, $l_2 \in C_2 \setminus B$. Furthermore, we have $C_1 \cap C_2 = \emptyset$ because otherwise B, C_1, C_2 and j_1, j_2, j_3 for some $j_3 \in C_1 \cap C_2$ would satisfy Condition (ii). So for each $i \in X'$ with $|\mathcal{M}(i)| \geq 3$ we have the situation depicted in the left of Figure 5.1.

If there now exists some other point $i' \in C_1$ with $|\mathcal{M}(i')| \geq 3$, then we have to be in the same situation for this point again if $i' \notin B$. In particular this implies also that $|\mathcal{M}(j)| \geq 3$ for some $j \in A$, so we can assume that $i' \in B$. We now repeat this process until we either get an element that we had before – implying that Condition (ii) holds – or we arrive at some set A that has exactly one $i \in A$ with $|\mathcal{M}(i')| \geq 3$.

Repeating the same process for C_2 instead of C_1, we finally arrive at the following situation: For some $\nu \in \mathbb{N}$ there exist $i_1, \ldots, i_\nu$ and $A_1, \ldots, A_\nu$ such that $\mathcal{M}(i_1) =$

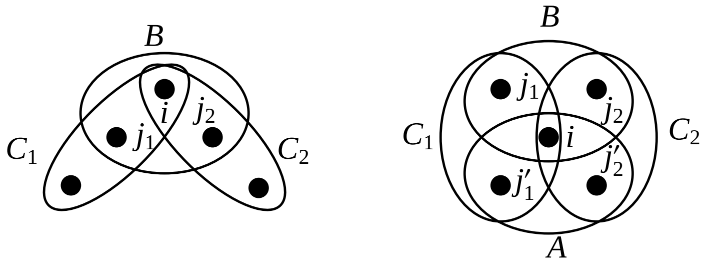


FIGURE 5.1. Situations for $i \in X'$ with $\mathcal{M}(i) = 3$ and $\mathcal{M}(i) \geq 4$, respectively.

$\{A_1, A_2\}$, $\mathcal{M}(i_l) = \{A_{l-1}, A_l, A_{l+1}\}$ for $1 < l < \nu$, $\mathcal{M}(i_\nu) = \{A_{\nu-1}, A_\nu\}$, and $A_l = \{i_{l-1}, i_l, i_{l+1}\}$ for $1 < l < \nu$.

We now consider two cases: First suppose $\nu \equiv 2 \bmod 3$. Then choose $0 < \epsilon < \min\{x_{i_l}, 1 - x_{i_l} \,|\, 1 \leq l \leq \nu\}$ and consider $x' \in \mathbb{R}^n$ defined by

$$x'_l = \begin{cases} x_l + \epsilon, & \text{if } l = i_m \text{ and } m \equiv 1 \bmod 3, \\ x_l - \epsilon, & \text{if } l = j_m,\ l = i_m \text{ and } m \equiv 1 \bmod 3, \\ x_l, & \text{else.} \end{cases}$$

Then $x \neq x'$ and $x \in I$, a contradiction. So suppose $\nu \not\equiv 2 \bmod 3$. Then it is easily seen that the values of x_{i_l} for $1 \leq l \leq \nu$ are determined by the values $\sum_{i \in A_1 \setminus \{i_1, i_2\}} x_i$ and $\sum_{i \in A_\nu \setminus \{i_{\nu-1}, i_\nu\}} x_i$. This implies that $\mathcal{M}'' := \mathcal{M}' \setminus \{A_l \,|\, 1 \leq l \leq \nu\}$ with

$$F' := F \cap \left\{ x \in \mathbb{R}^n \;\middle|\; x_{i_l} = \begin{cases} 1, & \text{if } l \equiv 1 \bmod 3 \\ 0, \text{else}, & \end{cases} \text{ for all } 1 \leq l \leq \nu \right\}$$

if $\nu \equiv 0 \bmod 3$ and

$$F' := F \cap \left\{ x \in \mathbb{R}^n \;\middle|\; x_{i_l} = \begin{cases} 1, & \text{if } l \equiv 1 \bmod 3 \\ 0, \text{else}, & \end{cases} \text{ for all } 1 \leq l < \nu \right\}$$

if $\nu \equiv 2 \bmod 3$ would also have been a valid choice at the beginning, but $\mathcal{M}'' \subsetneq \mathcal{M}'$ contradicts the minimality of $\mathcal{M}'$.

(vi): Suppose there exists some $i \in X'$ with $\mathcal{M}(i) \geq 4$. As in the proof of (v), we have to be in the situation depicted in the left of Figure 5.1 and there exists some $A \in \mathcal{M}(i) \setminus \{B, C_1, C_2\}$. Since Condition (i) cannot hold, every $j \in A$ has to be in some $C' \in \mathcal{M}(i)$ and, again by F2, there exist distinct $j'_1, j'_2 \in A$ and $C'_1, C'_2 \in \mathcal{M}(i)$ with $j'_1 \in C_1$, $j'_2 \in C'_2$ and $l'_1 \in C'_1 \setminus A$, $l'_2 \in C'_2 \setminus A$. Since Condition (i) cannot hold, we get $C'_1, C'_2 \in \{A, C_1, C_2\}$. However, if, for example, $C'_1 = C_1$ and $C'_2 = A$, then i, j'_1, j'_2 and A, B, C_1 would satisfy Condition (ii). Hence we have $C'_1 = C_1$ and $C'_2 = C_2$, or vice-versa. So we are in the situation depicted in the right of Figure 5.1. To obtain in addition some j with $\mathcal{M}(j) \leq 3$, there has to exist some $D \in \mathcal{M}'$ with $D \cap U \neq \emptyset$, where $U := A \cup B \cup C_1 \cup C_2$. Because Condition (i) cannot hold, we

get $|D \cap U| \geq 2$ and so F2 implies that either Condition (i) or Condition (ii) has to be satisfied, a contradiction. □

We will now show that under our assumptions at the beginning of the proof one of the Conditions (i) to (vi) has to be satisfied, which leads to a contradiction.

For each $A \in \mathcal{M}'$ we define $\tilde{A} := \{i \in A \cap X' \mid \mathcal{M}(i) \leq 2\}$. We have $\left|\tilde{A}\right| \geq 2$ for all $A \in \mathcal{M}'$, because otherwise we would have a situation satisfying one of Conditions (v) or (vi). Given some pair $(A, \delta) \in \mathcal{M}' \times X'$ with $\delta \in \tilde{A}$, we now give a way to construct a finite sequence $F(A, \delta) = (A_j, \alpha_j)_{1 \leq j \leq L(A,\delta)} \subset \mathcal{M}' \times X'$:

I $(A_1, \alpha_1) := (A, \delta)$.

II If there exists some $\gamma \in \tilde{A}_j$ such that $A_l \in \mathcal{M}(\gamma)$ for some $l < k$, then $L(A, \delta) = j$ and (A_j, α_j) is the last element of the sequence;

III else, if there exists some $\gamma \in \tilde{A}_j$ such that $\mathcal{M}(\gamma) = \{A_j, C\}$ for some $C \neq A_j$, then we set $A_{j+1} := C$ and $\alpha_{j+1} := \gamma$;

IV else, there exist a (unique) $\gamma \in \tilde{A}_j$ with $\mathcal{M}(\gamma) = \{A_j\}$; then $L(A, \delta) = j$ and (A_j, α_j) is the last element of the sequence.

The existence of the $\gamma \in \tilde{A}_j$ in Case IV follows from the fact that $\left|\tilde{A}_j\right| \geq 2$ and its uniqueness from F1. Obviously, $F(A, \delta)$ ends in either Case II or in Case IV. Suppose there exist some pair (A, δ) ending up in Case II. Then $\alpha_1, \ldots, \alpha_{L(A,\delta)}$ and $A_1, \ldots, A_{L(A,\delta)}$ obviously satisfy Condition (ii) if $L(A, \delta)$ is odd and Condition (iv) if μ_A is even – a contradiction. Hence for each starting pair $(A, \delta) \in \mathcal{M}' \times X'$ with $\delta \in \tilde{A}$ we end up in Case IV. The unique element γ occurring there will be denoted $f(A, \delta)$.

Now choose some $B \in \mathcal{M}'$. By F4 and $\left|\tilde{B}\right| \geq 2$ there exists some $\delta \in B$ with $|\mathcal{M}(\delta)| = 2$, say $\mathcal{M}(\delta) = \{B, C\}$ for some $C \neq B$. We now construct the sequences $F(B, \delta) = (B_j, \alpha_j)_{1 \leq j \leq L(B,\delta)}$ and $F(C, \delta) = (C_j, \gamma_j)_{1 \leq j \leq L(C,\delta)}$. Define

$$
\begin{aligned}
i_0 := f(B,\delta), \quad i_1 := \beta_{L(B,\delta)} & \qquad A_1 := B_{L(B,\delta)} \\
\ldots & \qquad \ldots \\
i_{L(B,\delta))} := \beta_1 = \delta = \gamma_1, & \qquad A_{L(B,\delta)} := B_1, A_{L(B,\delta)} := B_1, \\
\ldots & \qquad \ldots \\
i_{L(B,\delta)+L(C,\delta)-1} := \gamma_{L(C,\delta)}, i_{L(B,\delta)+L(C,\delta)} := f(B,\delta) & \qquad A_{L(B,\delta)+L(C,\delta)} := C_{L(C,\delta)} \,.
\end{aligned}
$$

Now if $e := L(B, \delta) + L(C, \delta)$ is odd, then these $i_0, \ldots, i_e$ and $A_1, \ldots, A_e$ satisfy Condition (iii). So e must be even.

Suppose there exists some $1 < j < e$ and some $\alpha \in A_j$ with $\alpha \neq i_{j-1}, i_j$. Then we distinguish two cases: First, assume that $\mathcal{M}(\alpha) = \{A_j\}$. Then either j is odd and $i_0, \ldots, i_{j-1}, \alpha$ and $A_1, \ldots, A_j$ satisfy Condition (iii), or j is even, hence $e - j + 1$ is odd and $\alpha, i_j \ldots, i_e$ and $A_j, \ldots, A_e$ satisfy Condition (iii). So assume that $D \in \mathcal{M}(\alpha)$ for some $D \neq A_j$. Now we construct the sequence $F(D, \alpha) = (D_j, \delta_j)_{1 \leq j \leq L(D,\alpha)}$. Then either $j + L(D, \alpha)$ is odd and $i_0, \ldots, i_{j-1}, \alpha = \delta_1, \ldots, \delta_{L(D,\alpha)}, f(D, \alpha)$ and $A_1, \ldots, A_j$, $D_1, \ldots, D_{L(D,\alpha)}$ satisfy Condition (iii) or $j + L(D, \alpha)$ is even, hence $e - j + L(D, \alpha) + 1$ is

odd and, similarly, $i_e \dots, i_j, \alpha = \delta_1, \dots, \delta_{L(D,\alpha)}$, $f(D,\alpha)$ and $A_e, \dots, A_j, D_1, \dots, D_{L(D,\alpha)}$ satisfy Condition (iii).

This shows that for each $\alpha \in A_j$ with $1 < j < e$ we have $\mathcal{M}(\alpha) = \{A_{j-1}, A_j\}$ or $\mathcal{M}(\alpha) = \{A_j, A_{j+1}\}$. By F1, this implies $\alpha = i_j$ or $\alpha = i_{j-1}$, respectively. Furthermore, it follows from this fact and the construction of $F(B,\delta)$ and $F(C,\delta)$ that $\alpha \in A_1 \setminus A_2$ implies $\alpha = i_0$ and $\alpha \in A_e \setminus A_{e-1}$ implies $\alpha = i_e$. Thus, each A_j, $1 \le j \le e$, has exactly two elements. Hence x has to satisfy the equations

$$x_{i_l} + x_{i_{l-1}} = 1, \quad \text{for all } 1 \le l \le e\,.$$

This implies that $x_{i_l} = x_{i_0}$ if l is odd and $x_{i_l} = 1 - x_{i_0}$ if l is even. In particular, $\sum_{l=0}^{e} x_{i_l} = x_0 + \sum_{l=1}^{e/2}(x_{i_l} + x_{i_{l-1}}) = e/2 + x_0$ is not an integer, hence there exists some $\gamma \in X' \setminus \{i_0, \dots, i_e\}$. We distinguish two cases: If $\mathcal{M}(\gamma) = \emptyset$, then choose some $0 < \epsilon < \min\{x_{i_0}, 1 - x_{i_0}, x_\gamma\}$ and define the point $x' \in \mathbb{R}^n$ via

$$x'_i = \begin{cases} x_i - \epsilon, & \text{if } i = i_l \text{ for some even } l, \\ x_i + \epsilon, & \text{if } i = \gamma, \text{ or } i = i_l \text{ for some odd } l \\ x_i, & \text{else.} \end{cases}$$

It is easily checked that $x' \in I$, a contradiction.

In the case $\mathcal{M}(\gamma) \neq \emptyset$ there exist some $B^\star \in \mathcal{M}'$ with $\gamma \in B^\star$. We can now argue as before: By F4 and $\left|\tilde{B^\star}\right| \ge 2$ there exists some $\delta^\star \in B^\star$ with $\left|\mathcal{M}(\delta^\star)\right| = 2$, say $\mathcal{M}(\delta^\star) = \{B^\star, C^\star\}$ for some $C^\star \neq B^\star$. This leads us to $i_0^\star, \dots i_{e^\star}^\star$ and $A_1^\star, \dots A_{e^\star}^\star$ having the same properties as $i_0, \dots, i_e$ and $A_1, \dots, A_e$. Choose some $0 < \epsilon < \min\{x_{i_0}, 1 - x_{i_0}, x_{i_0^\star}, 1 - x_{i_0^\star}\}$ and define the point $x' \in \mathbb{R}^n$ via

$$x'_i = \begin{cases} x_i - \epsilon, & \text{if } i = i_l \text{ for some even } l \text{ or } i = i_l^\star \text{ for some odd } l, \\ x_i + \epsilon, & \text{if } i = i_l \text{ for some odd } l \text{ or } i = i_l^\star \text{ for some even } l, \\ x_i, & \text{else.} \end{cases}$$

It is easily checked that $x' \in I$, our final contradiction. □

6. Compatibility and k-Weak Compatibility of splits of X

In this section, we present some corollaries of Theorem 1.2. Recall that two splits $\{A,B\}$ and $\{C,D\}$ are called *compatible* if one of the four intersections $A \cap C$, $A \cap D$, $B \cap C$, or $B \cap D$ is empty; a set $\mathcal{S}$ of splits is called *compatible* if each pair of elements of $\mathcal{S}$ is compatible (see e.g., [20]).

We first consider the case $k = 2$. In this case, for a split $\{A,B\}$ of X, the splits S_A and S_B of $\Delta(2,n)$ are clearly equal.

Corollary 6.1 (Corollary 6.3 and Proposition 6.4 in [13]). *Let $\mathcal{S}$ be a set of splits of X.*

(a) $\mathcal{S}$ is compatible if and only if $\mathcal{T} := \{S_A \text{ split of } \Delta(2,n) \,|\, A \in S, S \in \mathcal{S}\}$ is a compatible set of splits of $\Delta(2,n)$

(b) $\mathcal{S}$ is weakly compatible if and only it is 2-weakly compatible.

Proof. (a) Follows from Lemma 3.3.

(b) Condition (a) of Theorem 1.2 reduces exactly to the usual definition of weak compatibility of splits of X, since the condition on the cardinality is redundant for $k = 2$. Condition (c) can never occur if $k = 2$, and Condition (b) can only occur in the case $\nu = 1$. In this case, however, $i_0, i_3, i_1, i_2 \in X$ and the splits S_1, S_2, S_3 also fulfil Condition (a) for some $i_0 \in X \setminus (S_1(i_1) \cup S_2(i_2) \cup S_3(i_3))$.

□

Note that this last proof follows directly from the definition of weak compatibility for splits of sets and splits of polytopes, whereas the proof of [13, Proposition 6.4] uses the uniqueness of the split decomposition for metrics [1, Theorem 2] and weight functions for polytopes [13, Theorem 3.10].

We now consider the case $k \geq 3$.

Proposition 6.2. *Let $\{A, B\}$, $\{C, D\}$ be two distinct splits of X and $\mathcal{T} := \{S_F$ split of $\Delta(k, n) \mid F \in \{A, B, C, D\}\}$ be the set of corresponding splits of $\Delta(k, n)$. Then we have:*

(a) If $\mathcal{T}$ is compatible, then $\{A, B\}$ and $\{C, D\}$ are compatible.
(b) If $\{A, B\}$ and $\{C, D\}$ are compatible, then there exists at most one non-compatible pair of splits in $\mathcal{T}$.
(c) If $\{A, B\}$ and $\{C, D\}$ are compatible and $A \cap C = \emptyset$, then $\mathcal{T}$ is compatible if and only if $k = 2$ or $|A \cup C| \geq n - k + 2$.

Proof. (a) By Lemma 3.3, if $\{A, B\}$ and $\{C, D\}$ are not compatible, the only possibility for S_A and S_C or S_A and S_D to be compatible is that $|A \cup C| \geq n-k+1$ or $|A \cup D| \geq n-k+1$, respectively. However, since $D = X \setminus C$, these two conditions cannot be true at the same time.

(b),(c) We assume without loss of generality (for (b)) that $A \cap C = \emptyset$. By Lemma 3.3, it follows that S_A and S_B, S_B and S_D, S_B and S_D, S_B and S_C, and S_A and S_D are compatible, so it only remains to consider the pair S_A and S_C. For this pair of splits Lemma 3.3 implies that it is compatible if and only if $|A \cup C| \geq n - k + 2$ or $k = 2$.

□

Corollary 6.3. *Let $\mathcal{S}$ be a compatible set of splits of X. Then $\mathcal{S}$ is k-weakly compatible for all $k \geq 2$.*

Proof. This follows directly from Theorem 1.2: If either of the properties (a), (b), or (c) would hold, then, for example, the pair of splits $\{A_1, X \setminus A_1\}$ and $\{A_2, X \setminus A_2\}$ would not be compatible. □

We conclude by remarking that each of the three conditions in Theorem 1.2 become weaker as k increases:

Corollary 6.4. *Let $\mathcal{S}$ be a set of splits of X and $k \geq 3$. If $\mathcal{S}$ is k-weakly compatible, then it is l-weakly compatible for all $2 \leq l \leq k$. In particular, a k-weakly compatible set of splits is weakly compatible.*

7. k-Dissimilarity Maps from Trees

Let $T = (V, E, l)$ be a weighted tree consisting of a vertex set V, an edge set E and a function $l : E \to \mathbb{R}_{>0}$ assigning a weight to each edge. We assume that T does not have any vertices of degree two and that its leaves are labelled by the set X. Such trees are also called *phylogenetic trees*; see Figure 1.1 for an example and Semple and Steel [20] for more details. As explained in Figure 1.1, we can define a k-dissimilarity map D_T^k by assigning to each k-subset $K \subset X$ the total length of the induced subtree. Each edge $e \in E$ defines a split $S_e = \{A, B\}$ of X by taking as A the set of all leaves on one side of e and as B the set of leaves on the other. It is easily seen that

$$D_T^k = \sum_{e \in E} l(e) \delta_{S_e}^k . \tag{7.1}$$

We now show how this decomposition of D_T^k is related to its split decomposition.

Proposition 7.1. *Let D be a k-dissimilarity map on X with $|X| \geq 2k - 1$. Then $D = D_T^k$ for some tree T if and only if $\mathcal{S}_D$ is compatible and $D_0 = 0$ in the split decomposition of D. Moreover, if this holds, then the tree T is unique.*

Proof. Suppose the split decomposition of D is given by

$$D = \sum_{S \in \mathcal{S}} \alpha_S^D \delta_S^k$$

for some compatible set $\mathcal{S}$ of splits of X. Then Equation (7.1) shows that for the tree T whose edges correspond to the splits in $S \in \mathcal{S}$ with weights α_S^D we have $D_T^k = D$.

Conversely, if $D = D_T^k$ for some weighted tree, Equation (7.1) is a decomposition of D_T^k. By Corollary 6.3, this decomposition is coherent and the uniqueness part of Theorem 1.1 completes the proof. □

This gives us a new proof of the following Theorem by Pachter and Speyer:

Theorem 7.2 ([19])**.** *Let T be a weighted tree with leaves labelled by X and no vertices of degree two, and $k \geq 2$. If $|X| \geq 2k - 1$, then T can be recovered from D_T^k.*

Proof. Compute the split decomposition of D. The proof of Proposition 7.1 now shows how to construct a tree T' with $D = D_{T'}^k$, and the uniqueness part of this proposition shows that $T = T'$. □

8. Remarks and Open Questions

8.1. **Tight-Spans.** It was shown in [13, Proposition 2.3] that the set of inner faces of a regular subdivision $\Sigma_w(P)$ of a polytope P is anti-isomorphic to a certain realisable polytopal complex, the *tight-span* $\mathcal{T}_w(P)$ of w with respect to P. If $P = \Delta(2,n)$ and $w_d := -d$ for a metric d on X then $\mathcal{T}_{w_d}(\Delta(2,n))$ is the tight-span T_d of the metric space (X,d); see Isbell [16] and Dress [7]. In particular, if d is a tree metric, then T_d is isomorphic to that tree. For a k-dissimilarity map D one can similarly consider the tight-span $\mathcal{T}_{w_D}(\Delta(k,n))$. However, Proposition 6.2 shows that $\mathcal{T}_{w_D}(\Delta(k,n))$ is not necessarily a tree for $k \geq 3$. As an example, we depict in Figure 8.1 the tight-span $\mathcal{T}_{w_{D_T^3}}(\Delta(3,6))$ where T is the tree from Figure 1.1. Even though it is not a tree, note that the non-trivial splits corresponding to the edges of T can be easily recovered from $\mathcal{T}_{w_{D_T^3}}(\Delta(3,6))$. It would be interesting to understand better the relationship between the structure of $\mathcal{T}_{w_D}(\Delta(k,n))$ and the split decomposition of D in case D has no split-prime component.

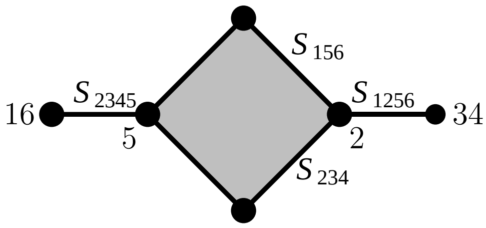


FIGURE 8.1. The tight-span of the subdivision of $\Delta(3,6)$ induced by the 3-dissimilarity map D_T^3 coming from the tree T in Figure 1.1. Note, that the three non-trivial splits $\{16,2345\}$, $\{34,1256\}$, (corresponding to the splits S_{2345}, S_{1256} of $\Delta(3,6)$, respectively) and $\{156,234\}$ (corresponding to the two splits S_{156}, S_{234} of $\Delta(3,6)$) can be recovered from the tight-span, as indicated in the figure.

8.2. **Matroid Subdivisions, Tropical Geometry, and Valuated Matroids.** A subdivision Σ of $\Delta(k,n)$ is called a *matroid subdivision* if all 1-dimensional cells $E \in \Sigma$ are edges of $\Delta(k,n)$, or, equivalently, if all elements of Σ are matroid polytopes. The space of all weight functions w inducing matroid subdivisions is called the *Dressian.* The elements of the Dressian correspond to (uniform) valuated matroids (see [12, Remark 2.4]) and to tropical Plücker vectors (see Speyer [21, Proposition 2.2]). The corresponding weight function w then defines a so called *matroid subdivision* of $\Delta(k,n)$. The *tropical Grassmannian* (see [22]) is a subset of the Dressian. It was shown by Iriarte [15] with methods developed by Bocci and Cools [2], and Cools [3] that for a weighted tree T, the weight function $w_{D_T^k}$ is a point in the tropical Grassmannian and hence in the Dressian. Corollary 6.3 now implies that $w_{D_T^k}$ is indeed in the interior of the cone of the Dressian spanned by

the split weights $w^k_{S_e}$ for all splits S_e corresponding to edges e of T. In the language of matroid subdivisions this implies that starting from a compatible set $\mathcal{S}$ of splits of X the set $\{S_A \text{ split of } \Delta(k,n) \mid A \in S, S \in \mathcal{S}\}$ of splits of $\Delta(k,n)$ induces a matroid subdivision. Establishing that other sets of splits satisfying the requirements of Theorem 5.1 also have this property could lead to a further understanding of the Dressian.

8.3. **Computation of the Split Decomposition and Tree Testing.** In [19], Speyer and Pachter raise the question how to test whether a given k-dissimilarity map D on X comes from a tree. Our results suggest the following simple algorithm: Compute the split indices α^D_S for all splits of X, test whether $D_0 = 0$ in the split decomposition (1.1), and whether the split system $\mathcal{S}_D$ is compatible. Equation (2) in [13] gives an explicit formula for the indices $\alpha^{w_D}_{w_{S_A}}$ and hence for the split indices α^D_S, however this involves the computation of the tight-span $\mathcal{T}_{w_D}(\Delta(k,n))$ whose number of vertices can be in general exponential in n. It would be interesting to derive a simpler formula for the split indices similar to the one existing in the case $k = 2$ given by Bandelt and Dress [1, Page 50]. This might yield a polynomial algorithm to test whether a given k-dissimilarity map D on X comes from a tree.

References

[1] H.-J. Bandelt A. W. M. Dress, A canonical decomposition theory for metrics on a finite set, Adv. Math. 92 (1992), 47–105.

[2] C. Bocci F. Cools, A tropical interpretation of m-dissimilarity maps, Appl. Math. Comput. 212 (2009), 349–356.

[3] F. Cools, On the relation between weighted trees and tropical Grassmannians, J. Symbolic Comput. 44 (2009), 1079–1086.

[4] J. A. De Loera, J. Rambau, F. Santos, Triangulations: Structures for algorithms and applications, Springer, 2010.

[5] M. M. Deza I. G. Rosenberg, n-semimetrics, European J. Combin. 21 (2000), 797–806.

[6] M. M. Deza M. Laurent, Geometry of cuts and metrics, Springer-Verlag, Berlin, 1997.

[7] A. W. M. Dress, Trees, tight extensions of metric spaces, and the cohomological dimension of certain groups: a note on combinatorial properties of metric spaces, Adv. in Math. 53 (1984), 321–402.

[8] A. M. Gabrièlov, I. M. Gel′fand, M. V. Losik, Combinatorial computation of characteristic classes, Functional Anal. Appl. 9 (1975), 48–49.

[9] C. Hayashi, Two dimensional quatification based on the measure of dissimilarity among three elements, 24 (1972), 251–257.

[10] W. J. Heiser M. Bennani, Triadic distance models: axiomatization and least squares representation, J. Math. Psych. 41 (1997), 189–206.

[11] S. Herrmann, K. T. Huber, V. Moulton, A. Spillner, Recognizing treelike k-dissimilarities, 2011, submitted.

[12] S. Herrmann, A. Jensen, M. Joswig, B. Sturmfels, How to draw tropical planes, Electron. J. Combin. 16 (2009), Research Paper 6, 26 p.

[13] S. Herrmann M. Joswig, Splitting polytopes, Münster J. Math. 1 (2008), 109–141.

[14] D. H. Huson D. Bryant, Application of phylogenetic networks in evolutionary studies., Mol. Biol. Evol. 23 (2006), 254–267.

[15] B. Iriarte Giraldo, Dissimilarity vectors of trees are contained in the tropical Grassmannian, Electron. J. Combin. 17 (2010), Note 6, 7 p.

[16] J. R. Isbell, Six theorems about injective metric spaces, Comment. Math. Helv. 39 (1964), 65–76.

[17] S. Joly G. Le Calvé, Three-way distances, J. Classification 12 (1995), 191–205.

[18] D. Levy, R. Yoshida, L. Pachter, Beyond pairwise distances: Neighbor-joining with phylogenetic diversity estimates, Mol. Biol. Evol. 23 (2006), 491–498.

[19] L. Pachter D. E. Speyer, Reconstructing trees from subtree weights, Appl. Math. Lett. 17 (2004), 615–621.

[20] C. Semple M. Steel, Phylogenetics, Oxford University Press, Oxford, 2003.

[21] D. E. Speyer, Tropical linear spaces, SIAM J. Discrete Math. 22 (2008), 1527–1558.

[22] D. E. Speyer B. Sturmfels, The tropical Grassmannian, Adv. Geom. 4 (2004), 389–411.

[23] M. J. Warrens, n-way metrics, J. Classification 27 (2010), 173–190.

[24] G. M. Ziegler, Lectures on polytopes, Springer-Verlag, New York, 1995.

School of Computing Sciences, University of East Anglia, Norwich, NR4 7TJ, UK

E-mail address: `s.herrmann@uea.ac.uk, vincent.moulton@cmp.uea.ac.uk`